



Multipathing Configuration for Software iSCSI Using Port Binding

TECHNICAL WHITE PAPER

Table of Contents

Multipathing for Software iSCSI	3
Configuring vmknic-Based iSCSI Multipathing	3
a) Configuring the Network	3
1. Using GUI	3
2. Using CLI	6
b) Enabling the iSCSI Software Adapter	7
1. Using GUI	7
2. Using CLI	9
c) Activating vmknic-Based Multipathing for Software iSCSI	9
1. Configuring Port Binding on vSphere 5.0 Using GUI	9
2. Configuring Port Binding on vSphere 5.0 Using CLI	12
3. Configuring Port Binding on vSphere 4.x Using CLI	12
iSCSI Re-login	12
Verifying Multipath Configuration	13
1. Verifying the iSCSI LUN Configuration	13
2. Verifying vmknic-Based Multipath Failover	13
3. Verifying Load Balancing Using the Round-Robin Policy	14
Conclusion	15
About the Author	15

Multipathing for Software iSCSI

Multipathing between a server and storage array provides the ability to load-balance between paths when all paths are present and to handle failures of a path at any point between the server and the storage. Multipathing is a de facto standard for most Fibre Channel SAN environments. In most software iSCSI environments, multipathing is possible at the VMkernel network adapter level, but not the default configuration.

In a VMware vSphere® environment, the default iSCSI configuration for VMware® ESXi™ servers creates only one path from the software iSCSI adapter (vmhba) to each iSCSI target. To enable failover at the path level and to load-balance I/O traffic between paths, the administrator must configure port binding to create multiple paths between the software iSCSI adapters on ESXi servers and the storage array.

Without port binding, all iSCSI LUNs will be detected using a single path per target. By default, ESX will use only one vmknic as egress port to connect to each target, and you will be unable to use path failover or to load-balance I/O between different paths to the iSCSI LUNs. This is true even if you have configured network adapter teaming using more than one uplink for the VMkernel port group used for iSCSI. In case of simple network adapter teaming, traffic will be redirected at the network layer to the second network adapter during connectivity failure through the first network card, but failover at the path level will not be possible, nor will load balancing between multiple paths.

Some of the user guides and documentation refer to vmknic-based software iSCSI multipathing as “port binding” or simply as “software iSCSI multipathing.” This paper provides an overview of how to enable vmknic-based software iSCSI multipathing, as well as the procedure by which to verify port binding configuration.

Configuring vmknic-Based iSCSI Multipathing

To enable vmknic-based software iSCSI multipathing, you must:

- Create two VMkernel port groups and connect one uplink to each of them.
- Bind each VMkernel network adapter to the software iSCSI adapter. Then run a rediscovery of iSCSI targets to detect multiple paths to them.

Refer to the following sections to learn the network configuration and port binding procedure required to enable vmknic-based multipathing for software iSCSI.

a) Configuring the Network

Enabling multipathing removes the ability to route to storage. The storage and the VMkernel port must have an IP address in the same network.

In the vSphere 5.0 release, you can use GUI wizard or CLI commands to configure your network; in vSphere 4.x, you can use only CLI commands.

1. Using GUI

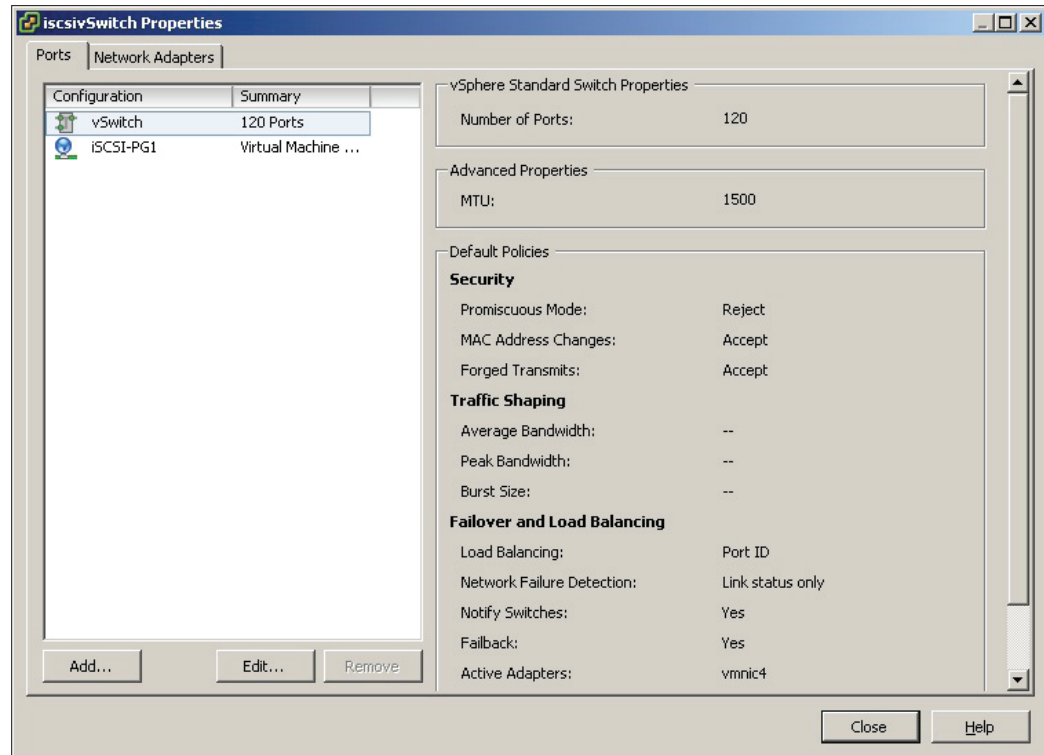
To enable vmknic-based multipathing for software iSCSI, it is recommended that you have two or more physical network adapters for iSCSI on your host machine.

- 1) Connect to the ESXi server using the VMware vSphere® Client™
- 2) Click the **Configuration** tab > **Networking**.

- 3) Click **Add Networking**.
- 4) Select **VMkernel** and click **Next**.
- 5) Select **Create a vSphere standard switch** to create a new vSwitch.
- 6) Select all the adapters you want to configure for iSCSI traffic and click **Next**.
- 7) Provide a **Network Label** for the VMkernel port group and click **Next**.
- 8) Specify the IP address settings and click **Next**.
- 9) Review the details on the **Summary** screen and click **Finish**.

Next, to create an additional VMkernel port group to configure port binding, complete the following steps:

- 1) Click the **Properties** tab of the vSwitch you created for iSCSI.
- 2) Create additional VMkernel ports for all network adapters that you connected by selecting the **Add** tab under the **Ports** tab in the **vSwitch Properties** dialog box.

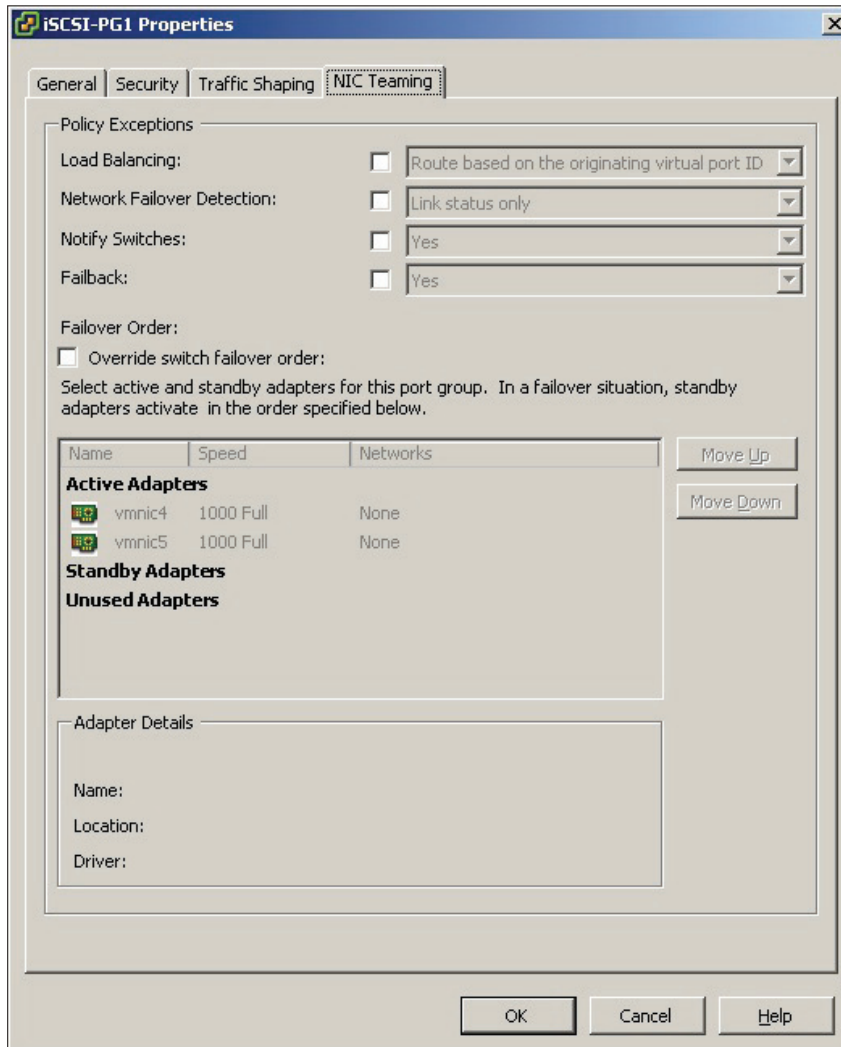


- 3) Select **VMkernel** and click **Next**.
- 4) Provide a **Network Label** for the VMkernel port group and click **Next**.
- 5) Specify the IP address settings and click **Next**.
- 6) Review the details on the **Summary** screen and click **Finish**.

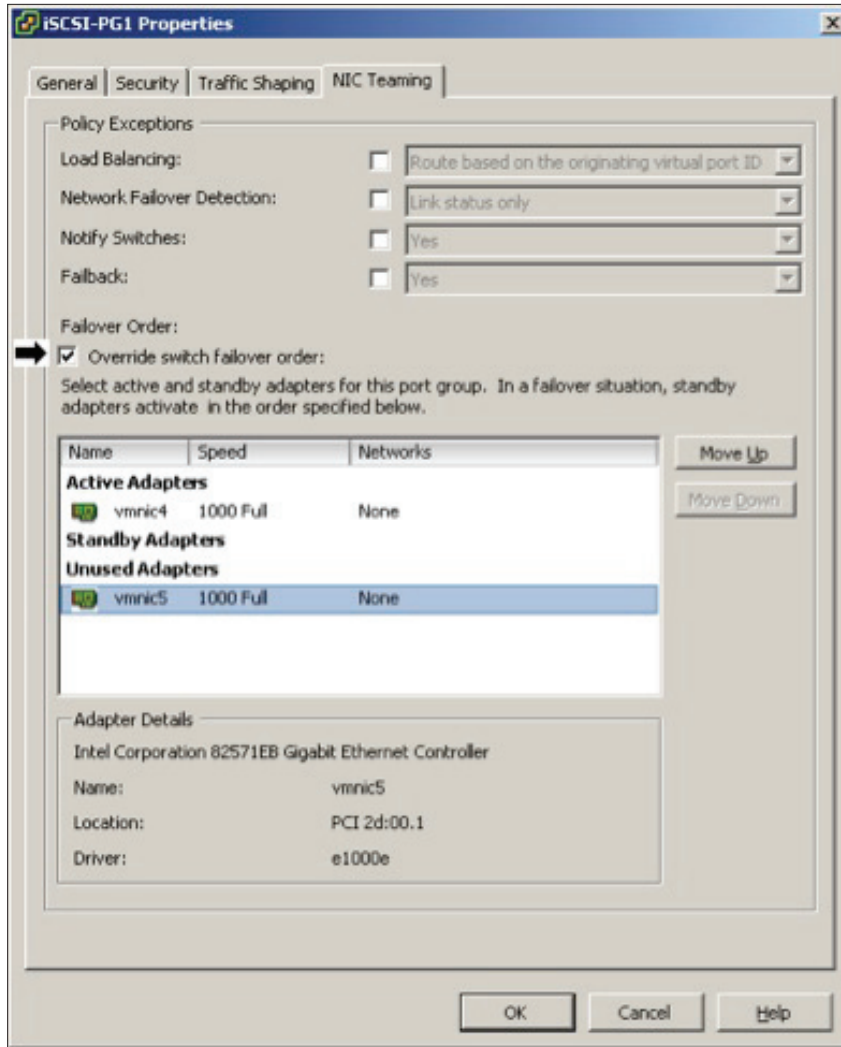
You can also create VMkernel port groups on different vSwitches. However, if your VMkernel network adapters are on the same subnet, they must be configured on the same vSwitch.

All network adapters that you added to the vSwitch will appear as active for each VMkernel port on the vSwitch. Ensure that you override this configuration so that each VMkernel port maps to only one active adapter.

- 1) In the **Ports** tab of the **vSwitch Properties** dialog box, select a VMkernel port and click **Edit**.
- 2) Click the **NIC Teaming** tab and check **Override switch failover order**.



- 3) Keep only one adapter under **Active Adapters** and use **Move Down** to move other adapters under **Unused Adapters**.



- Repeat steps 1-3 for each VMkernel port on the vSwitch, ensuring that each port has its own unique active adapter.

2. Using CLI

Run the following commands to create the vSwitch and VMkernel port groups, as well as to map each VMkernel port to one active adapter.

- Create the vSwitch: **esxcfg-vswitch -a iscsivSwitch.**
- Add the first uplink to the vSwitch: **esxcfg-vswitch -L vmnic4 iscsivSwitch.**
- Add the first port group on the vSwitch: **esxcfg-vswitch -A iSCSI-PG1 iscsivSwitch.**
- Add the second uplink to the vSwitch: **esxcfg-vswitch -L vmnic5 iscsivSwitch.**
- Add the second port group on the vSwitch: **esxcfg-vswitch -A iSCSI-PG2 iscsivSwitch.**

- 6) Map each VMkernel port to use only one active adapter and to move the second network adapter to the unused adapter list: `esxcfg-vswitch -N vmnic4 -p iSCSI-PG1 iscsivSwitch.`
- 7) Map each VMkernel port to use only one active adapter and to move the second network adapter to the unused adapter list: `esxcfg-vswitch -N vmnic5 -p iSCSI-PG2 iscsivSwitch.`
- 8) Assign an IP address to the first port group: `esxcfg-vmknic -a -i 192.168.100.1 -n 255.255.255.0 iSCSI-PG1.`
- 9) Assign an IP address to the second port group: `esxcfg-vmknic -a -i 192.168.100.2 -n 255.255.255.0 iSCSI-PG2.`

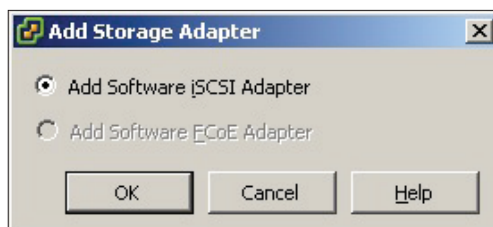
b) Enabling the iSCSI Software Adapter

1. Using GUI

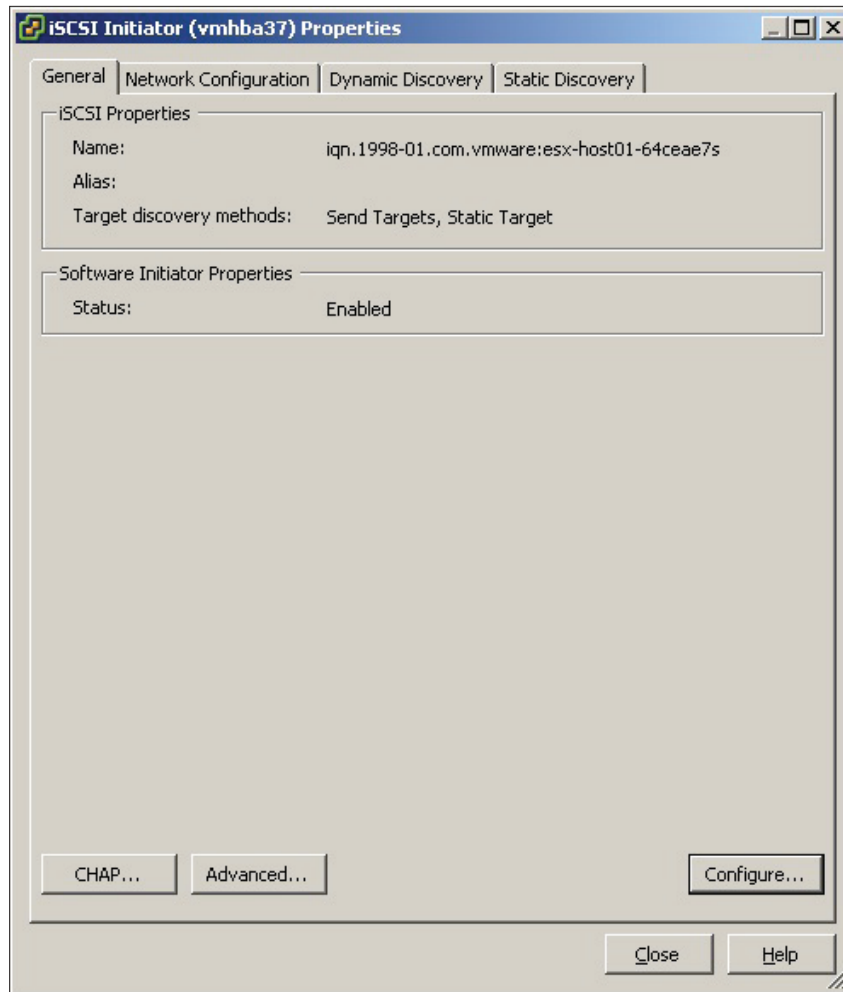
To access iSCSI targets, you must enable the software iSCSI initiator on the VMware ESXi server.

Procedure

- 1) Connect to the ESXi server using vSphere Client.
- 2) Click **Configuration > Storage Adapters**.
- 3) Enable the iSCSI software adapter.
 - a. For ESXi server releases prior to vSphere 5.0, **iSCSI Software Adapter** is listed under the **Storage Adapters** tab.
 - i. Select **iSCSI Software Adapter** and click **Properties**.
 - ii. Under the **General** tab, click **Configure**.
 - iii. Check **Enabled** to enable the initiator.
 - iv. To change the default iSCSI initiator name, enter **iSCSI initiator name** in the **iSCSI Name** box.
 - b. For vSphere 5.0, you must add **iSCSI Software Adapter** to the **Storage Adapters** category.
 - i. Click **Configuration > Storage Adapters**.
 - ii. Click **Add** and check **Add Software iSCSI Adapter**.



- iii. To change the default iSCSI initiator name, enter **iSCSI initiator name** in the **iSCSI Name** box under the **Configure** tab in the **iSCSI Software Adapter Properties** window.



- 4) Enter the iSCSI target address in **Static/Dynamic Discovery**.

Dynamic Discovery – Specify the addresses for *Send Targets* discovery. The iSCSI initiator sends a *Send Targets* request to each of the specified addresses, and the discovered targets are added to the static discovery list.

Static Discovery – A list of IP addresses and iSCSI names of targets to connect to. This list can be filled in based on dynamic *Send Targets* requests or entered individually. This is a list of targets with which ESX attempts to establish sessions.

- 5) Click **Close** to finish iSCSI initiator configuration.

2. Using CLI

Run the following commands:

- 1) Enable the software iSCSI: **esxcfg-swiscsi -e**
- 2) To change the default iSCSI initiator name, set the initiator iqn:
- **esxcli iscsi adapter set --name iqn.1998-01.com.vmware:esx-host01-64ceae7s -A vmhbaXX**
- 3) Add the iSCSI target discovery address:
- **esxcli iscsi adapter discovery sendtarget add -a 192.168.100.13:3260 -A vmhbaXX**

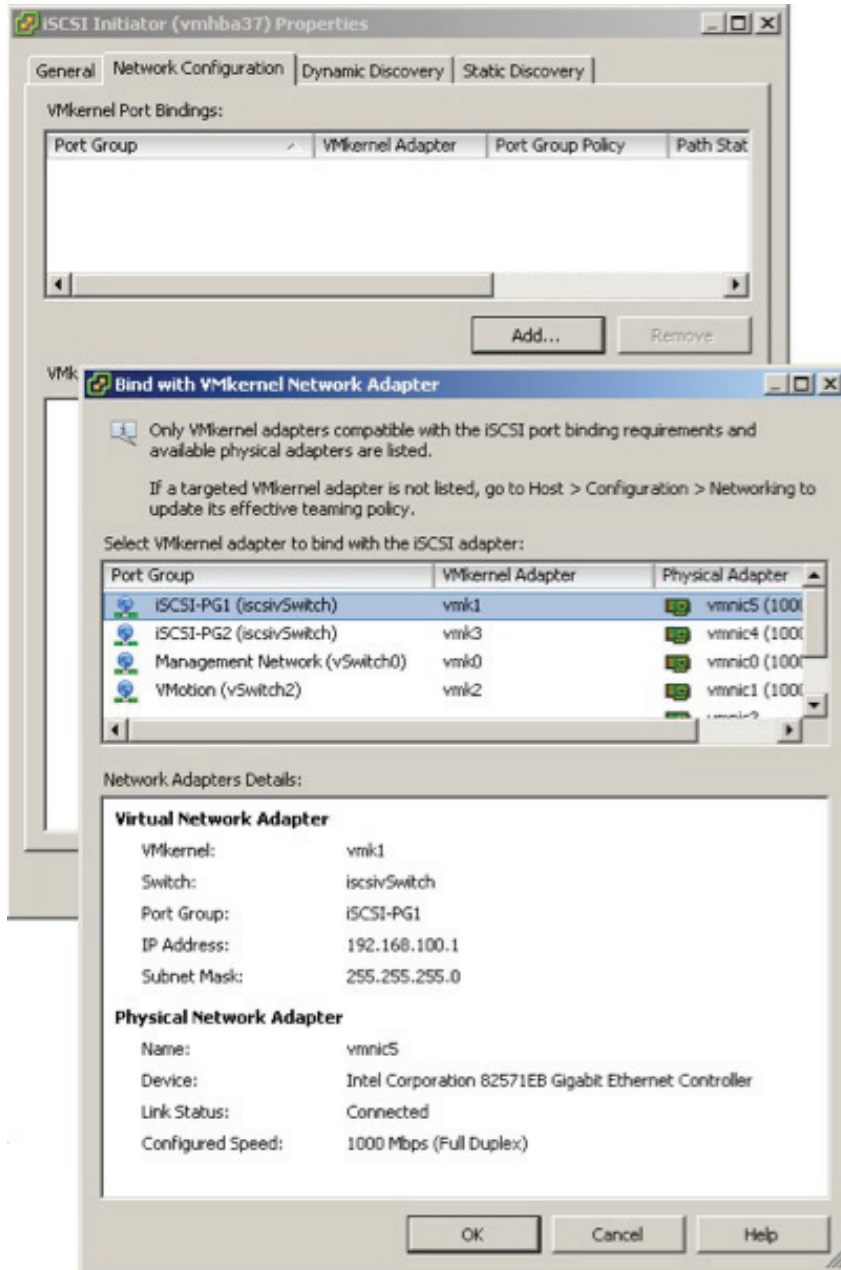
*NOTE: vmhbaXX is the software iSCSI adapter vmhba ID. To find the vmhba ID, you can use the command **esxcfg-scsidevs -a** and check for the vmhba ID of "iSCSI software adapter."*

c) Activating vmknick-Based Multipathing for Software iSCSI

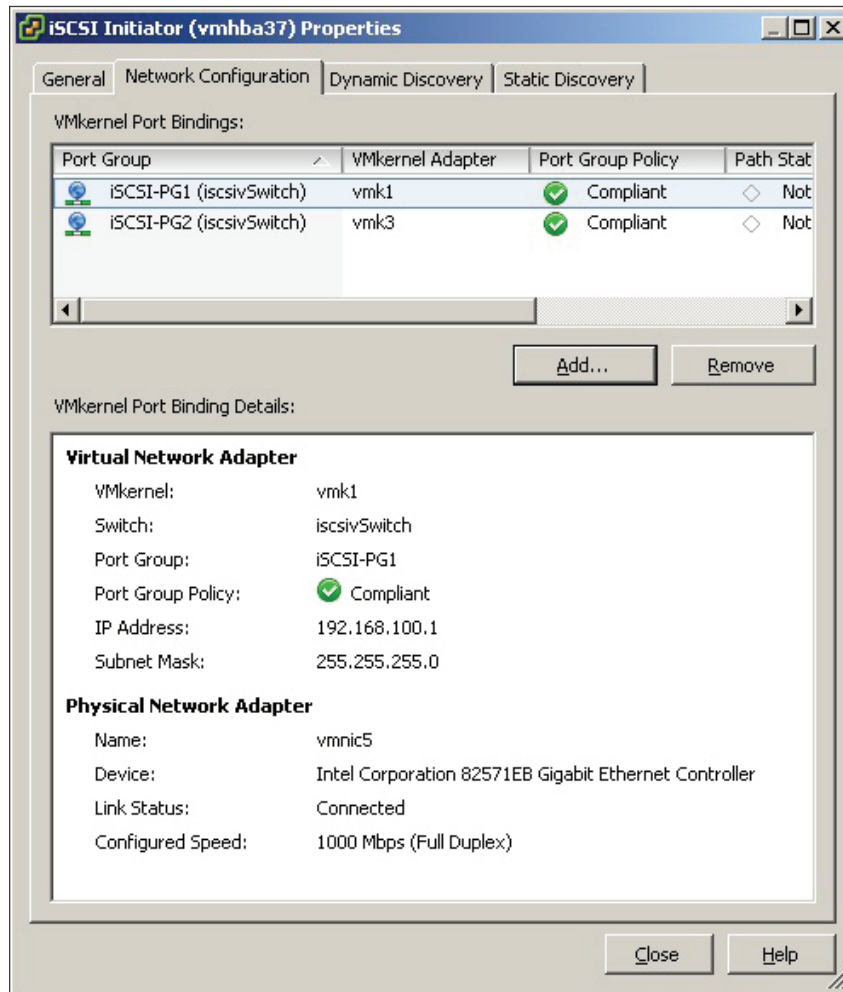
VMware vSphere 5.0 has added a new UI interface to support multipathing configuration for the software iSCSI adapter using port binding. (Earlier ESX/ESXi releases had only the command-line interface (CLI) option to configure port binding.)

1. Configuring Port Binding on vSphere 5.0 Using GUI

- a) Connect to the ESXi server using vSphere Client.
- b) Click the **Configuration** tab and select **Storage Adapters**.
- c) Select **iSCSI Software Adapter** and click **Properties**.
- d) Click the **Network Configuration** tab and click **Add** to bind the VMkernel network adapter to the software iSCSI adapter.
- e) The bind with the VMkernel adapter window is displayed, listing all the VMkernel adapters compatible with iSCSI port binding requirements. Select the VMkernel network adapter you want to bind to the software iSCSI adapter and click **OK**.



f) Repeat steps d) and e) until you bind all the required VMkernel adapters to the iSCSI adapter.



NOTE: The port group policy status might report as noncompliant for the VMkernel network adapter for the following reasons:

- The VMkernel network adapter is not connected to an active physical network adapter or it is connected to more than one physical network adapter.
 - The VMkernel network adapter is connected to standby physical network adapters.
 - The active physical network adapter got changed.
- g) Close the **iSCSI Initiator Properties** window.
- h) Select the software iSCSI adapter and run **Rescan** to verify that multiple paths are available for iSCSI LUNs.

2. Configuring Port Binding on vSphere 5.0 Using CLI

Run the following commands from the vSphere 5.0 host CLI to bind the VMkernel network adapter to the iSCSI adapter:

```
- esxcli iscsi networkportal add --nic vmk0 --adapter vmhbaXX  
- esxcli iscsi networkportal add --nic vmk1 --adapter vmhbaXX
```

Where **vmk0** and **vmk1** are the VMkernel network adapters compatible with iSCSI port binding requirements. You can get the list of VMkernel network adapters connected to all vSwitches by using the command **esxcfg-vmknic -l. vmhbaXX** is the software iSCSI adapter vmhba ID.

You can verify the binding details by using the following command:

```
esxcli iscsi networkportal list --adapter vmhbaXX.
```

Then rescan the software iSCSI adapter to verify that multiple paths are available for iSCSI LUNs.

3. Configuring Port Binding on vSphere 4.x Using CLI

To enable vmknic-based multipathing for software iSCSI, run the following CLI commands:

```
- esxcli swiscsi nic add --nic vmk0 --adapter vmhbaXX  
- esxcli swiscsi nic add --nic vmk1 --adapter vmhbaXX
```

Where **vmk0** and **vmk1** are the VMkernel network adapters compatible with iSCSI port binding requirements. You can get the list of VMkernel network adapters connected to all vSwitches by using the command **esxcfg-vmknic -l. vmhbaXX** is the software iSCSI adapter vmhba ID.

After configuring port binding, you can verify the binding details by using the following command:

```
esxcli swiscsi nic list --adapter vmhbaXX.
```

Then rescan the software iSCSI adapter to verify that multiple paths are available for iSCSI LUNs.

iSCSI Re-login

If you have an already established iSCSI session before port binding configuration, you can remove the existing iSCSI sessions and log in again for the port binding configuration to take effect.

To list the existing iSCSI sessions, run the following command:

```
- esxcli iscsi session list --adapter vmhbaXX
```

To remove existing iSCSI sessions, run the following command:

```
- esxcli iscsi session remove --adapter vmhbaXX
```

To enable sessions as per the current iSCSI configuration, run the following command:

```
- esxcli iscsi session add --adapter vmhbaXX
```

Verifying Multipath Configuration

1. Verifying the iSCSI LUN Configuration

After you complete the port binding configuration and perform the iSCSI adapter rescan, verify that iSCSI LUNs are detected and displayed with the available paths. For example, if you have bound two VMkernel network adapters to the iSCSI adapter, you will see two paths to each LUN connected to the iSCSI targets. To confirm, run the **esxcfg-mpath** command from CLI or click the **Paths** tab for LUNs on the GUI.

Sample output

Before port binding, there is one path available to each target (T0, T1).

```
~ # esxcfg-mpath -bd naa.600601602af129007d698975a6f2e011
naa.600601602af129007d698975a6f2e011 : DGC iSCSI Disk (naa.600601602af129007d698975a6f2e011)
  vmhba37:CO:T0:LO LUN:0 state:active iscsi Adapter: iqn.1998-01.com.vmware:prme-cpdqe14-swiscsi
  vmhba37:CO:T1:LO LUN:0 state:active iscsi Adapter: iqn.1998-01.com.vmware:prme-cpdqe14-swiscsi
~ #
```

After port binding two VMkernel network adapters to the iSCSI adapter, there are two paths available to each target (T0, T1).

```
~ # esxcfg-mpath -bd naa.600601602af129007d698975a6f2e011
naa.600601602af129007d698975a6f2e011 : DGC iSCSI Disk (naa.600601602af129007d698975a6f2e011)
  vmhba37:CO:T0:LO LUN:0 state:active iscsi Adapter: iqn.1998-01.com.vmware:prme-cpdqe14-swiscsi
  vmhba37:C1:T0:LO LUN:0 state:active iscsi Adapter: iqn.1998-01.com.vmware:prme-cpdqe14-swiscsi
  vmhba37:CO:T1:LO LUN:0 state:active iscsi Adapter: iqn.1998-01.com.vmware:prme-cpdqe14-swiscsi
  vmhba37:C1:T1:LO LUN:0 state:active iscsi Adapter: iqn.1998-01.com.vmware:prme-cpdqe14-swiscsi
```

2. Verifying vmknick-Based Multipath Failover

To identify the VMkernel network adapter currently used by the iSCSI adapter to transmit I/O, run the **esxtop** command. In the following example, I/O is transmitted through vmk1 (vmnic5).

USED-BY	TEAM-PNIC	DNAME	PKTTX/s	MbTX/s	PKTRX/s	MbRX/s	%DRPTX	%DRPRX
Management	n/a	vSwitch0	0.00	0.00	0.00	0.00	0.00	0.00
vmnic0	-	vSwitch0	2.17	0.01	4.53	0.01	0.00	0.00
vmk0	vmnic0	vSwitch0	2.17	0.01	3.15	0.01	0.00	0.00
iSCSI-200G	vmnic0	vSwitch0	0.00	0.00	0.00	0.00	0.00	0.00
Management	n/a	iscsiVSwitch	0.00	0.00	0.00	0.00	0.00	0.00
vmnic4	-	iscsiVSwitch	0.79	0.00	3.54	0.01	0.00	0.00
vmnic5	-	iscsiVSwitch	72.65	0.11	145.30	1.38	0.00	0.00
vmk1	vmnic5	iscsiVSwitch	72.65	0.11	143.14	1.37	0.00	0.00
vmk3	vmnic4	iscsiVSwitch	0.79	0.00	1.58	0.00	0.00	0.00
Management	n/a	vSwitch2	0.00	0.00	0.00	0.00	0.00	0.00
vmnic1	-	vSwitch2	0.00	0.00	0.20	0.00	0.00	0.00
vmk2	vmnic1	vSwitch2	0.00	0.00	0.00	0.00	0.00	0.00

To verify vmknick-based multipath failover, disable any one of the network adapters connected to the VMkernel port group used for iSCSI. Verify that the path status is updated correctly and that the paths connected to the failed network adapter card are marked as dead.

```

~ # esxcli storage nmp satp set -s naa.600601602af129007d698975a6f2e011
naa.600601602af129007d698975a6f2e011 : DGC iSCSI Disk (naa.600601602af129007d698975a6f2e011)
  vmhba37:CO:TO:LO LUN:0 state:dead iscsi Adapter: Unavailable Target: Unavailable
  vmhba37:C1:TO:LO LUN:0 state:active iscsi Adapter: iqn.1998-01.com.vmware:prme-cpdqe14-swiscsi
  vmhba37:CO:T1:LO LUN:0 state:dead iscsi Adapter: Unavailable Target: Unavailable
  vmhba37:C1:T1:LO LUN:0 state:active iscsi Adapter: iqn.1998-01.com.vmware:prme-cpdqe14-swiscsi
~ #

```

In the **esxtop** output, the failed network adapter is marked as **fallback** and the I/O is transferred using the other VMkernel port group.

USED-BY	TEAM-PNIC	DNAME	PKT/s	Mb/s	PKT/s	Mb/s	%DRPTX	%DRPRX
Management	n/a	vSwitch0	0.00	0.00	0.00	0.00	0.00	0.00
vmnic0	-	vSwitch0	6.69	0.02	10.83	0.01	0.00	0.00
vmk0	vmnic0	vSwitch0	6.69	0.02	8.86	0.01	0.00	0.00
iSCSI-2006	vmnic0	vSwitch0	0.00	0.00	0.00	0.00	0.00	0.00
Management	n/a	iscsivSwitch	0.00	0.00	0.00	0.00	0.00	0.00
vmnic4	-	iscsivSwitch	1153.56	0.64	2290.20	25.18	0.00	0.00
vmk3	vmnic4	iscsivSwitch	1153.56	0.64	2288.42	25.18	0.00	0.00
vmk1	fallback	iscsivSwitch	0.20	0.00	0.20	0.00	0.00	0.00
vmnic5	-	iscsivSwitch	0.00	0.00	0.00	0.00	100.00	0.00
Management	n/a	vSwitch2	0.00	0.00	0.00	0.00	0.00	0.00
vmnic1	-	vSwitch2	0.00	0.00	0.00	0.00	0.00	0.00
vmk2	vmnic1	vSwitch2	0.00	0.00	0.00	0.00	0.00	0.00

3. Verifying Load Balancing Using the Round-Robin Policy

After configuring multipath failover for the software iSCSI, you can balance I/O traffic between the paths using the round-robin path policy.

NOTE: Verify that the iSCSI array supports the round-robin path policy by contacting the array vendor.

Change the path selection policy (PSP) for iSCSI LUNs either from the command line, by selecting the **Manage Paths** option, or by running the following command:

```

- esxcli storage nmp satp set -s SATP_NAME -P VMW_PSP_RR

```

To find the SATP name used to claim the iSCSI LUNs, run the following command:

```

- esxcli storage nmp device list -d naa.600601602af129007d698975a6f2e011

```

Run the **esxtop** command to verify that I/O is distributed across the available paths for the iSCSI LUN.

To increase aggregated IO performance one can change RoundRobin IO Operation Limit to a lower value from default 1000.

USED-BY	TEAM-PNIC	DNAME	PKT/s	Mb/s	PKT/s	Mb/s	%DRPTX	%DRPRX
Management	n/a	vSwitch0	0.00	0.00	0.00	0.00	0.00	0.00
vmnic0	-	vSwitch0	4.33	0.06	8.27	0.01	0.00	0.00
vmk0	vmnic0	vSwitch0	4.33	0.06	6.50	0.01	0.00	0.00
iSCSI-2006	vmnic0	vSwitch0	0.00	0.00	0.00	0.00	0.00	0.00
Management	n/a	iscsivSwitch	0.00	0.00	0.00	0.00	0.00	0.00
vmnic4	-	iscsivSwitch	458.35	0.60	930.69	8.23	0.00	0.00
vmk3	vmnic4	iscsivSwitch	458.35	0.60	927.34	8.21	0.00	0.00
vmk1	vmnic5	iscsivSwitch	767.86	0.41	1518.00	16.62	0.00	0.00
vmnic5	-	iscsivSwitch	767.86	0.41	1519.58	16.62	0.00	0.00
Management	n/a	vSwitch2	0.00	0.00	0.00	0.00	0.00	0.00
vmnic1	-	vSwitch2	0.00	0.00	0.00	0.00	0.00	0.00
vmk2	vmnic1	vSwitch2	0.00	0.00	0.00	0.00	0.00	0.00

Conclusion

Configuring vmknic-based multipathing for the software iSCSI adapter will help vSphere users enable failover at the path level as well as balance I/O traffic between the paths.

About the Author

Sudhish P T is a Staff Engineer of the VMware R&D group responsible for validation of storage on VMware vSphere releases. His area of expertise is virtualization, with a focus on vSphere storage technologies. He has 12 years of industry experience and has been working with VMware India, Bangalore, since December 2007.

